群智感知中基于社交属性及有效用户计算的任务分发机制 *

杨玉仁[1]，张书奎[1,2]，龙 浩[1]，张 力[1]

(1. 苏州大学 计算机科学与技术学院, 江苏 苏州 215006; 2. 江苏省无线传感网高技术研究重点实验室, 南京 210003)

摘 要：随着无线传感器网络的高速发展和多种移动智能设备的普及，移动群智感知(mobile crowd sensing，MCS)成为移动计算的核心。利用群智感知可完成大规模、复杂环境及社会感知任务，其中任务分发是这种应用中的一个重要环节。针对任务分发过程中存在感知环境复杂、用户数量达不到要求、收集数据质量低等问题，提出一种基于社交属性及有效用户计算的任务分发机制(effective user calculation，EUC)，该机制具有根据任务来筛选用户的特点，从用户角度看，EUC 考虑了用户的社会性，由用户的社交网络传递相关信息来增加平台的有效用户数；从平台的角度看，EUC 可根据任务的接收和提交情况，动态调整有效用户的积分，从而保障整个系统的有效用户数。理论分析和实验结果表明，所提出的机制可提高系统的任务分发效率，并改善了收集数据的质量。

关键词：群智感知；任务分发；有效用户；社交属性

中图分类号：TP393.07 doi: 10.3969/j.issn.1001-3695.2018.01.0002

Task distribution mechanism based on social attribute and effective user calculation in crowd sensing

Yang Yuren[1], Zhang Shukui[1, 2], Long Hao[1], Zhang Li[1]

(1. *School of Computer Science & Technology*, *Soochow University*, *Suzhou Jiangsu* 215006, *China*; 2. *Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks*, *Nanjing* 210003, *China*)

Abstract: Mobile Crowd Sensing (MCS) is the core of mobile computing with the rapid development of wireless sensor networks and the popularity of mobile intelligent devices. The use of crowd sensing can be completed large-scale, complex environment and social awareness tasks, which the task distribution is an important part of this application. In order to solve the problem that the perceived environment is complex, the number of users can not meet the requirements and the quality of the collected data is low. we propose a task distribution mechanism (EUC) based on social attribute and effective user calculation. According to the task to filter the characteristics of the user, from the user point of view, EUC consider the user's social, the user's social network through the relevant information to increase the number of effective users; from the platform point of view, EUC according to the task of receiving and Submit the situation, dynamically adjust the effective user's points, thus ensuring the effective number of users of the entire system. Theoretical analysis and experimental results show that the mechanism proposed in this paper can improve the task distribution efficiency of the system and improve the quality of the collected data.

Key words: crowd sensing; task distribution; effective users; social attributes

0 引言

随着移动通信和感知技术的迅速发展，涌现出大量创新性的应用和服务，特别是智能手机的普及，将本文迅速带入物联网时代（Internet of things，IoT)，使得对现实世界更全面、更大规模地感知各种物理对象和环境状况成为可能[1,2]。移动感知和移动互联网技术的进步触发了人们对日常携带设备(如智能手机、可穿戴设备等)或者车载传感器(如 GPS)的研究来监测那些不能被单一节点独立检测到的大规模现象。大量诸如此类的移动终端组成了一个无处不在的、具有强大感知能力的移动无线传感器网络，它们内在的流动性使一种新的快速增长的感知模式成为可能，可随时随地地感知人及周边环境信息，这些设备可以通过移动互联网实现有意识或者无意识的合作，用以进行任务分发、数据收集、数据处理等操作,使得人们可以及时地获取自

基金项目：国家自然科学基金资助项目（61201212）；江苏省自然科学基金资助项目（BK2011376）；江苏省“六大人才高峰”资助项目（2014-WLW-010）；苏州市融合通信重点实验室资助项目（SKLCC2013XX）；江苏省产学研前瞻性项目（BY2012114）；软件新技术与产业化协同创新中心资助项目

作者简介：杨玉仁（1993-），男（通信作者），安徽人，硕士研究生，主要研究方向为无线传感器网络（aeyanr0718@163.com）；张书奎（1966-），男，内蒙古人，教授，博导，博士，主要研究方向为无线传感器网络、分布式计算、智能信息处理、并行和分布式系统等；龙浩（1984-），男，湖南人，讲师，硕士，主要研究方向为移动群体计算、分布式计算等；张力（1980-），男，安徽人，讲师，硕士，主要研究方向为群智计算、物联网、无线网络等.

身及其所处环境的状况，如城市环境监测[3]、智能电网[4]、交通状况[5]、医疗保健[6]等，它们大大扩展了人类感知世界的维度，改变了人们感知世界的方式，开启了移动互联的新领域-移动群智感知和计算大门。

由于群智计算中的节点是用户个人拥有，会跟随人们的日常生活经常移动，具有动态的、缺乏稳定的网络拓扑结构的特点。目前，许多研究关注节点的移动，试图根据某种特征建立相应的移动模型，从而可以采用一些规则预测节点所在的位置。可是，用户的移动是社会性的，预测节点的移动轨迹是一件非常困难和复杂的事情，在这种环境下很难将传感任务分发到指定的节点。同时，这种网络以数据为中心，并不关心数据是来自哪些节点，但是在目前的研究中，用户须提供相应的个人信息才能获得感知任务，例如：位置、时间、活动轨迹等一些列敏感信息，这些信息综合起来可以准确标志某个用户，这使得用户的相关隐私被暴露[7]，因此，为了大规模使用群智感知应用，用户的隐私保护是至关重要的。

群智感知的任务分发，具有代表性的是 Bubble-sensing 模式[8]，该模式将任务信息登记在相关服务器上，移动到该区域的用户广播任务信息，接收任务后进行数据采集，为了保持存在任务传播者，该机制设置了一些静态性质的用户，持续传播任务信息。但该模式主要存在以下几个问题：(1)过分依赖用户间的相遇机会进行任务分发，由于用户的社会特征及其移动轨迹的不确定，用户不能及时获取任务信息，降低了任务分发速度；(2)用户间转发任务具有盲目性，导致任务的拒绝率比较高，而且也需要尊重用户的意愿，是否要接受任务推送；(3)分发任务时没有进行用户的区分，造成恶意用户提交低质量数据，导致污染最终的分析结果；(4)没有完善的奖励机制来促进用户的参与。

为了解决上述问题，本文提出了一种基于有效用户计算的任务分发机制。重点在于提高任务分发过程的效率，主要贡献如下：

a)加入了用户社交元素，使得用户接收到任务后通过社交网络通知满足条件的朋友加入系统，扩展了任务分发的方式，减少了系统开销，提高了系统的参与人数。

b)使用差分隐私机制使用户相关信息的泛化，确保用户隐私安全。

c)提出一种反映有效用户状态的积分机制，根据用户的完成任务情况，自动更新用户的有效积分，作为执行下次任务时，筛选用户的依据，同时也可以屏蔽一些恶意用户的污染行为。

d)提出一种有效的激励机制，促进用户的参与。

1　相关研究

在当前的群智感知系统研究中，存在许多针对用户隐私保护、数据可信度、数据安全、奖励机制和数据传输效率等方向的研究。

典型的有文献[8]所提出的 Bubble-sensing 群智感知框架，该架构中把任务分发给一部分用户后，依赖于用户的移动性，转发给与他相遇的用户，这些用户可能在感知区域也可能不在，由于参与用户的社会性，其移动轨迹不确定，使得部分用户不能及时获取任务信息，无法在规定时间内完成任务，降低了任务的转发速度，甚至造成用户上传非感知地区的伪数据，如图 1 所示。

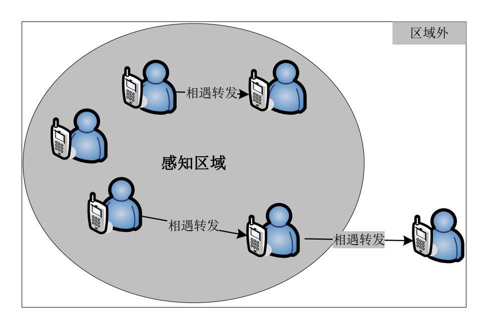


图 1　Bubble-sensing 模式任务转发策略

在这种模式下的任务转发机制，主要有两种，一是用户经过感知区域，由感知区域的服务器进行任务推送。另一种是依赖用户的移动性，使得部分用户携带任务信息，从而进行任务转发，其转发过程主要包含四个阶段：任务分发、数据接收、数据处理和数据应用，这四个阶段依次进行，其中任务分发阶段的主要功能是感知应用平台将任务信息发送给指定区域内的能够获取感知数据的用户，再由这些用户将任务分发给随机遇到的用户。其框架如图 2 所示。可以看出，在该机制中，任务分发依赖于用户的移动性，而系统将任务信息分发给一部分用户，要靠这些用户的移动来传播任务，显然，这样做会造成任务时间过长而且数据质量得不到保障，使得有些人用非感知区域的伪数据来代替监测区域的数据上传到服务器，从而污染整个数据库。

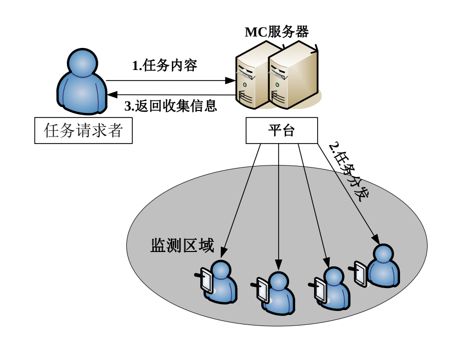


图 2　Bubble-sensing 模式的任务分发架构

文献[9]设计了一种旨在保护移动传感系统中的数据请求者和数据提供者的隐私保护平台，该平台使用云上的移动代理，根据用户的位置隐私偏好，映射移动用户到外部世界。这种结构使得用户可以从感兴趣的地理区域寻找数据。数据请求者在整个过程中都匿名，而且这些数据提供者证明他们是平台的合法用户，不论数据请求者还是数据提供者都有着良好的隐私保护，但是没有考虑到数据转换的效率，所提出的隐私保护方案，需要消耗大量资源来换取隐私安全，所以该方案在实施过程中

需要大量资源来支撑。

在文献[10]中，作者针对 MCS 系统提出了三个度量标准：作用、隐私和效率，并在三者之间寻求权衡，提出了一个灵活的优化框架，使得 MCS 平台可以和用户共同寻找上述三者之间的平衡点。由于上述问题的优化是 NP 难问题，该文献作者提出了几个优化算法来解决。合理的解决任务分配以及用户的隐私保护问题，而且做到了任务的高效推荐，这是该文献的亮点，但是并未考虑在数据传输过程中，数据的可信度问题，重点放在了如何权衡上述三者之间的方案讨论。

文献[11]讨论了感知数据量很大情况下的场景，考虑到请求者对获得的数据进行聚集和分析成为了一个困难问题的时候，尤其是当请求者是一个普通的智能手机用户或者是有着有限的存储和计算资源设备时。作者为MCS 系统引入了第三方架构，在工作者和请求者之间实现云，以缓解资源有限请求者的存储与计算负担，同时也考虑了身份隐私和数据隐私的保护。通过该方案，请求者可以验证云端计算结果的正确性。通过利用云计算来解决庞大数据的分析与检索，但是该文没有讨论目前存在不可信云的问题。

文献[12]也是借助云对数据进行分析和集合操作，当请求者进行集合操作时，用户的数据隐私和身份隐私得到很好保护。此外，请求者可以验证集合操作结果的正确性。作者还扩展提出的方案以支持数据预处理，在数据分析之前可以排除无效数据，增加了数据可信度。通过使用批量验证和数据更新方法，所提出的方案降低了计算成本。同时也解决了不可信云的问题。

文献[13]提出了一种关于建立在信息融合基础上的隐私保护方案，利用假名机制去克服参与者密度低的漏洞。为了进一步处理匿名攻击，基于 Paillier 密码系统和非交互式零知识验证下，促进和改善解决方案框架，也涵盖了用户责任问题，实现了一个全面的 MCS 基础设施，在复杂场景中也可充分保护到用户的位置隐私，文献[14]中也有着类似的做法。文献[15]中提出了一种隐私保护架构，可以在真实的移动设备全面实施，具有较高的效率，实用性和可扩展性，促进了可信赖的 MCS 应用的安全部署。文献[16]中，对参与环境感知的用户提供激励，通过适度的付款，以增加参与用户的数量，通过应用需求和供应变化的关系，决定感知数据的价值。分析感知频率、传感点远近和用户的偏好等影响他们的感知动机的因素。研究表明在传输距离少即有适当报酬的情况下，人们会积极地参与感知任务。

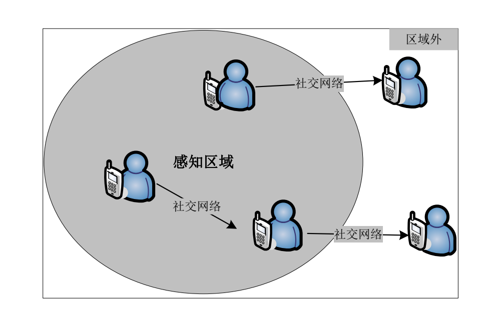
上述文献都是针对 MCS 中较为普遍的问题进行了一系列的探讨，并提出了相应的解决框架。本文主要关心的是群智感知系统中任务分发的问题，目前存在两种任务分发模式，直接转发与间接转发，直接转发是采取一对多的方式由平台向监测区域内的用户分发信息。文献[17,18]就是这种模式提出了相应的任务分发机制，但各自缺点也很明显，要么没有考虑到用户的情景信息，要么对 MCS 服务器的性能要求极高。间接转发模式是采取一对少的方式，服务器将任务信息先分发给少数的代理用户，然后再由代理用户来转发扩散。文献[19]基于这个模式提出了一些任务分发机制，但是在任务分发准确性上以及分发开销之间没有取得很好地平衡。

本文提出的任务分发模式，介于直接模式与间接模式之间，即半间接分发模式，即任务只能由系统发布给注册用户，由注册用户通过社交网络转发给其他的用户，这样可以保证用户在接受到任务后，执行任务的可靠性更高。

2　EUC 框架结构

本文所提出的任务转发方法，充分考虑了用户的社交属性来提高转发速度，如图 3 所示。用户在完成指定任务并提交后可以利用自身的社交属性，通过社交平台转发任务相关的信息，加快了传播速度，有较好的及时性。

将上述任务分发与数据收集分为以下几个阶段：用户注册、泛化注册信息、发布任务、注册任务、筛选用户、指定区域数据收集和数据粗糙过滤。上述几个部分每完成一次，更新有效用户的积分，并对效用户动态排名，其框架结构如图 4 所示。



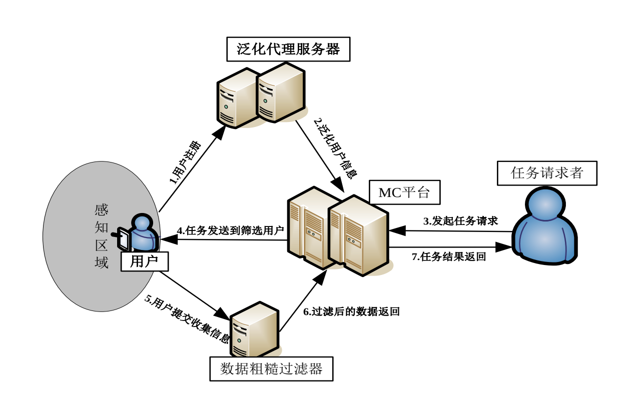


图 4　EUC 框架的任务分发模式

a）用户注册。由于 Bubble-sensing 机制不能很好解决用户的隐私保护问题，考虑在感知任务执行前，用户需要自身将信息注册到系统所认同的第三方代理，需要提交唯一标志信息，例如手机号，身份证等。第三方代理运用差分隐私机制[20,21]，使得用户是否向系统服务器提交信息的事实不会改变该用户在服务器的关联关系，这样，恶意用户无法从多次任务执行中来追踪或者推断出其他用户的身份信息。

b）泛化注册信息。将用户的信息根据各种指标进行泛化分类，生成一个毫无关联性的用户列表，分类完成后用户列表转交给服务器。

c）发布任务。大部分群智感知应用都是任务请求者拥有自己的服务进行任务分发，但本文中提出的是服务器与任务请求者分开，这样可以降低任务请求者的执行任务成本，确切说，MCS 服务器也相当于一个第三方代理，任务请求者把任务内容发送给 MCS 服务器即可，每个任务都包含用户列表。

d）注册任务。用户信息泛化注册到 MCS 平台，当用户登陆到平台后，首先获取地理位置，平台根据其地理位置以及有效积分来选择是否推送任务。如果系统给用户推送任务消息，用户会被注册到任务的用户列表中。

e）筛选用户。这个阶段，MCS 服务器根据历史表现计算用户的有效积分，并据此筛选任务分发给合适的用户，这里用户的积分根据其执行任务的表现动态更新。

f）数据收集。为了防止一些恶意用户利用非感知区域的数据来代替感知任务所需的信息，需要对用户的地理位置进行监测，用户一旦离开监测区域，要求立刻停止数据采集，而且监测区域外的用户无法获取感知任务信息，杜绝用户以假乱真的行为。

g）数据过滤。用户收集的数据上传到 MCS 平台的一个中间数据处理服务器，即粗糙数据过滤器，过滤掉明显的错误信息，并对提交这些数据的用户进行标记，再把用户与数据提交给 MCS 平台处理，从一定程度上保证了数据质量，也减轻 MCS 服务器的压力。

3 问题描述与相关定义

3.1 问题描述

群智感知网络的任务分发机制解决的主要问题是感知平台如何将任务信息分发给指定监测区域内能收集感知数据的移动用户，即有足够的用户去上传可靠的数据，使得任务高质量的完成，在研究中本文发现主要存在以下问题需要解决：

a）用户在移动过程中，用户的移动具有随机性，仅依赖用户的移动性进行任务信息分发，很大程度上限制了任务的传播能力，更重要的是，这种转发方式不但影响分发速度而且未考虑用户执行任务的效率。

b）需要对用户执行任务的效果进行动态评估，否则，平台中可能会存在部分恶意用户，这些用户的感知数据影响了分析结果的可信度，还造成了一些不必要的开销。

为解决上述问题，提出一种半间接的任务分发模式，这种模式基于本文提出的一种有效用户计算机制 EUC(effective user calculation)，参与任务分发的用户依靠自己的社交网络，将任务信息有目的性地发送给自己熟悉并且符合条件的用户，减少了分发时间，关键是提高了任务分发的质量，根据当前用户完成的情况对有效用户进行动态评估，使得在下次任务中筛选到合适的用户。

3.2 相关定义

下面定义所提算法中出现的一些概念。设某监测地区有任务 T 至少需要 N_1 个用户来完成，完成该任务后可获得酬劳为 P_t，给出的一些定义如下：

定义 1 有效积分 Q，用户获取任务的评价依据。初始情况下，新注册用户的有效积分都为 Q_1，后期根据用户完成任务情况动态更新，排名越靠前，被分发到任务的的可能性就越高。

定义 2 提交接受率 SAR（submit and accept rate），当用户向系统上传数据并被接受，由 SAR 来描述该过程。设当前某任务需要 N_1 个用户接收感知任务，最后只有 N_2 个用户的数据可以被 MCS 服务器接受，则

$$\mathrm{SAR} = N_2 / N_1 \tag{1}$$

定义 3 动态酬劳 R（reward），在执行任务过程中，消耗了用户的部分资源，不是所有人都愿意无条件付出，适当的奖励机制，可促进用户参与活动。基于上述的提交接收率，用户获得的酬劳定义为

$$\mathrm{R} = P_t \times SAR \tag{2}$$

定义 4 时间窗，为了保证平台获取数据的及时性，为每个任务都设定一个时间窗口进行限制，即$[S_i, E_i]$，S_i 表示任务的开始时间，E_i 表示任务结束时间。

定义 5 任务列表 TL（Task List）：任务分发根据地理位置进行，每个有效用户在某地理位置都有一个这样的列表，其主要属性有用户标志、时间点、位置信息、任务性质等。

定义 6 冷采集，用户在某个位置首次执行任务，TL 表为空，此过程为冷采集。

定义 7 热采集，用户在该位置历史上执行过感知任务，TL 表不为空，该过程为热采集。

从上述定义可以看出，SAR 可能会超过 1，即任务要求 N_1 个用户，却有超过 N_1 个用户去执行，因此，设定 SAR 最大为 1，超过 1 按照 1 来计算。可以看出如果大于 1，说明执行任务的用户已经超过了预期设定人数，在数据最后的处理上，数据的可靠性会有所提高而且有更多的选择，可以做到优取差弃。这样 SAR 值也可以看做数据质量纯度的近似反映。用户所获得的酬劳与用户完成任务直接关联，运用这一机制，可来促进用户使用社交平台来通知更多的人参与活动，但是为了保证任务发起者适当的成本，奖励设有任务的数量上限，因此 SAR 最大为 1，可以保证用户收入为 P_t，从而控制运营成本。针对一些实时性很强的任务，通过时间窗口对任务完成的实时性进行控制，例如，某时段某路口交通的实时情况，对时间要求就很高，需要设定时间窗口。

4 社交属性

本文核心思想就是将用户的社交属性纳入到群智感知中，可充分利用每个用户的朋友关系，实现任务信息的快速传达，从而使感知任务被按时完成。下面对用户的社交属性进行分析，如图 5 所示。

在图 5 中，用户 A_1 接收到任务信息，通过社交网络把任务信息转发出去，符合任务条件的用户 A_2、A_3、A_5、A_6 收到该信息，然后接收到任务的新这些用户利用社交网络再一层

一层的传递下去，形成一个依赖朋友关系的传播网络。当用户接收到任务并完成后，系统会分发给用户一个任务邀请码，邀请码的使用次数有限，这个限制可以根据具体任务来自确定。不可以让用户无限制的转发任务信息的原因，是为了防止一些恶意用户集聚一起，污染上传数据。

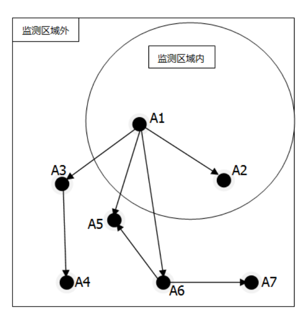


图 5　社交属性

对用户社交属性这种传播关系的描述，本文称之为“社交关联度”。在图 5 中，A_1 的出度为 4，A_1 的社交关联度量为 4。A_5 没有出度，只有来自 A_6 与 A_1 的入度，由于每个用户执行任务都有一个邀请码，如果 A_5 使用了 A_1 的邀请码，则 A_6 的这条边无效。从图中可知 A_6 的出度为 2，入度为 1，A_6 的社交关联度量为 1。表 1 显示图 5 中用户的社交关联度的值。

表 1　用户的社交关联度表

用户	出度	入度	关联度
A_1	4	0	4
A_2	0	1	0
A_3	1	1	1
A_4	0	1	0
A_5	0	2	0
A_6	2	1	1
A_7	0	1	0

邀请码的使用次数有限制，具体值由相应的任务来决定，可使用生成一个 10 以内的随机数来实现。用户可根据这个随机数有选择的传播任务信息，但传播次数不可以超过该限制数。

5　解决方案与算法

在本节中，本文详细解释所提出的算法，首选考虑基于有效用户筛选的贪婪算法。初始情况下，设 MCS 系统接收到任务发起者的任务 *T*，需要 N_1 个用户来执行感知任务，新注册用户 *U* 的积分为 Q_l，任务列表为 TL，存在少量用户。所设计筛选用户的贪婪算法，可使哪些用户添加到用户列表 *L* 中，算法如下：

Algorithm 1. Greedy Algorithm for select User

Input: T, Q_1, TL, N_1, U

Output : *L*　//用户列表

//　initialization

1: MCS 平台接收到任务首先作如下操作;

2: *L*<—Ø;

3: **if**　TL== Ø **then**

4:　说明该位置未执行过任务，所有的用户均可加入到用户列表 *L*;

5: else

6:　说明该地理位置使用过，再做如下判断;

7:　　**if** TL.Count > N_l **then** //TL.Count 表示上次执行用户数字

8:　把 TL 表中前 N_l 个用户直接加入到 L 表中，剩下的 TL.Count - N_l 个用户不加入到 *L* 表中;

9:　　**else**

10:　　把 TL 表积分高于 Q_l 的用户全部加入到 *L* 表中，去掉信用低用户;

11:　　**endif**

12: end if

13: **while** $Q == Q_l$ **do**

14:　*L* <—*U;* //在时间窗内监听新注册进来的用户 U

15: end while

16: TL = *L*　　//重新覆盖 TL 表，作为下次筛选依据

17: return *L*

通过上述提出的算法可以看出，MCS 平台在接收到任务请求者请求发起一个任务的时候，MCS 平台首先会检查该地区的任务列表 *TL* 是否为空，即该地区是否执行过感知任务，其中 *TL* 表存放的是该地区历史上执行过感知任务的有效用户。若 *TL* 表不为空，说明存在老用户，部分老用户是可以直接推送任务的。然后再判断 *TL* 表中用户数量与任务需求用户数量 N_l 相比较，当 *TL* 表中用户数量小于任务需求用户数量 N_l 时，把 *TL* 表中用户全部加入到可以执行任务的用户列表 *L* 中。否则，根据一个新注册的用户有效积分 Q_l 来作为阈值，进行新老用户筛选。当每次有新用户加入时，会根据新加入用户有效积分进行任务用户筛选，做到及时更新任务列表 *TL* 与用户列表 *L*。而对于那些有效积分不够的用户，对于本次任务则无权参加。考虑到任务发起者的成本问题，执行任务的用户数必须有个上限。

筛选完用户后，把相应的任务发送给筛选出的用户，然后再根据用户完成任务情况，动态更新用户积分，提出了一个有效用户排序算法（dynamic ranking algorithm for effective user，DREU）算法来解决这个问题。

当某地区需要执行的任务 *T* 完成后，返回用户任务列表 TL，表中表明 N_1 个用户接受任务，执行任务并提交的有 N_2 个用户，设 Q′为有效积分增加片段，*L* 为临时存放最终排序好的用户积分情况，下面是对 TL 每个用户 *U* 进行有效积分更新的算法描述。

Algorithm 2. Dynamic Ranking Algorithm for Effective User

Input : T,Q, Q′,TL,N1,N2,U

Output :TL

// initialization

```
SAR = N1 / N2;
if SAR > 1 then SAR = 1;
U = TL—>first        //获取 TL 表中第一个用户
while U != null do
    if U 是否执行并提交任务 then
        U.Q = U.Q + Q′ * SAR;   L<—U;
    else
        U.Q = U.Q – Q′ * SAR;   L<—U;
  end if
U = U—>Next;   //获取下一个用户
end while
sort(L);              //基于有效积分对 L 进行排序
TL = L;               //排序好的 L 在此覆盖 TL 表
return TL;
```

通过上述提出的算法可以看出，用户有效积分动态计算，首先根据用户对任务完成情况下，计算提交接受率 SAR，然后开始遍历任务列表 *TL* 中的每一个用户，判断任务列表 *TL* 中用户 *U* 是否进行任务提交，如果进行任务提交，则对用户 U 有效积分进行增加。否则，进行扣减有效积分。有效积分增加或减少都是以提交接受率 SAR 与 Q′有效积分乘积为标准。从而使有效积分动态更新。

上述算法中，充分利用了提交接受率 SAR，用户积分主要根据每次任务的完成情况来实现更新，可以看出参与并执行完成任务的用户越多，SAR 越大，证明采集到的数据越多，这对任务发起者来说有更多的选择。

在遍历 TL 表时，根据每个用户的完成情况动态刷新，接收到任务并完成，有效积分自然增加，而接受了任务没执行，则浪费了服务器的资源，要扣除有效积分。当用户提交数据后，粗糙数据过滤器进行检查，判断用户提交的数据是否存在问题，例如，要检测某个路段的噪声，某用户上传了一个 200 分贝，甚至更大的数据，这明显是恶意用户，此时服务器，对该用户进行标记，该用户在下一轮任务中不被选择。上述算法也可以看出有效积分的增长与数据纯度百分比有直接关系，每次有效积分的更新要根据当前用户完成任务情况，可防止有效用户积分虚高。

6　理论分析

本文提出的 EUC 机制，用户利用自身社交网络对任务间接转发，从某种程度上来说，用户的社交属性在一定程度上包含了相遇转发。例如，在某感知区域内收集的信息，以社交活动的形式邀请经过的人参与，这类似于相遇转发。将 EUC 机制与文献[8]的 BSD 模式进行对比，通过理论分析说明本文方法的可行性，首先引入以下几个定义：

这其中引用了文献[22]中相关证明的方法。

a)拒绝转发率 P_r：用户之间相遇时，相遇的用户转发任务信息，但是接收到信息的用户，由于某些原因拒绝转发，称为拒绝转发率。

b)任务执行率 T_r：接收任务后，随着时间的推移，用户执行任务的意愿会发生变化，可能存在用户收到任务，但是没有足够的时间去执行。设任务时间窗[t_1,t_2]这个时间段完成时间为 T_1，采集信息时间为 T_2，则任务可以执行时间 $T_r = T_1 - T_2$，任务执行率 T_r 为

$$T_r = \frac{T-t}{T} \tag{3}$$

c)相遇率 P_m：用户在指定区域与其他人相遇的概率，设某区域的感知任务需要 N_1 个用户，某用户在经过感知区域相遇了 m_1 个用户，相遇率 P_m 为：

$$P_m = \frac{m_1}{N_1} \tag{4}$$

d)社交通知率 P_n：通过社交网络通知其他用户参与任务执行的概率，设某感知区域的感知任务需要 N_1 个用户去完成，通过社交网络通知 n_1 个用户参与，则社交通知率 P_n 为：

$$P_n = \frac{n_1}{N_1} \tag{5}$$

定理 1　在进行任务分发时，与 BSD 相比，EUC 具有更高的成功率。

证明　由上述定义可知，BSD 任务分发的成功率，主要取决于用户的移动性，其分发成功率可表示为：

$$P_{BSD}(t) = \frac{T-t}{T} \times \frac{m_1}{N_1} \times P_r \tag{6}$$

而本文所提出的 EUC 机制中，任务的转发，不仅涉及到相遇转发，还要考虑用户的社交属性，转发成功率可表示为：

$$P_{EUC}(t) = \frac{T-t}{T} \times \frac{m_1}{N_1} \times P_r + \frac{T-t}{T} \times \frac{n_1}{N_1} \tag{7}$$

可以看出 $P_{EUC}(t) > P_{BSD}(t)$，即在执行相同的任务以及在相同时间的条件下，本文提出的 EUC 机制分发任务具有较高的成功率。证毕。

定理 2　在相同用户数的情况下执行任务，与 BSD 相比，EUC 具有更快的速度将任务分发给用户。

证明：设某用户在时间 *t* 开始任务转发，根据上述定义从两个方面进行。

a)上述证明可知，BSD 任务分发的成功率：

$$P_{BSD}(t) = T_r \times P_r \times P_m \tag{8}$$

EUC 机制任务分发的成功率：

$$P_{BSD}(t) = T_r \times P_r \times P_m + T_r \times P_n \tag{9}$$

对上述两式积分，即为在相同时间内，可以获得两种机制某用户转发给其他用户的数量。分别记为 N_{BSD}、N_{EUC}。

$$N_{BSD} = \int_0^T P_{BSD} dt = \int_0^T (T_r \times P_r \times P_m) dt \tag{10}$$

$$N_{EUC} = \int_0^T P_{EUC} dt = \int_0^T (T_r \times P_r \times P_m + T_r \times P_n) dt \tag{11}$$

显然，$N_{EUC} > N_{BSD}$，即 EUC 机制中用户分发任务的能力

比 BSD 要好。

b) 由第五节可知，提到了一个社交属性量化的名词社交关联度，即在用户接收到任务并完成后，用户可以通过社交方式把任务信息分发出去，本文设置社交关联度最大为 10，所以可知，假设某任务有 n 个人完成，在最理想的情况下，通过社交属性能召集到 $10*n$ 个人。当然这个情况几乎很难达到，但是最差的情况为召集到 n 个人完成任务。所以克制一个任务完成人数 N 的范围为。

$$n \le N \le 10*n$$

可知，BSD 方式主要依赖用户的随机移动进行任务转发，而在 EUC 中涉及到相遇转发只为任务转发其中一小部分，即 BSD 方式召集到人数，可以看做 EUC 最差的情况下召集到的人数。可知与 BSD 相比，EUC 具有更快的速度将任务分发给用户。证毕。

定理 3　EUC 机制的任务分发，用户上传的数据与 BSD 相比，具有更高的可靠性。

证明　该定理可以从三个方面进行证明。如下：

a)与 BSD 相比，EUC 机制中加入了时间窗和区域监测两种属性。对时间窗来说，提高了任务对时间的要求，在完成任务的时间上，更加严谨，对完成任务的时间范围有强行限制。

而区域监测，要求用户必须在指定区域内完成相关任务，对用户位置进行实时监测，杜绝了恶意用户使用非监测区域内的数据。可见在 EUC 机制中，用户所上传的数据可以做到在时间上与位置上的准确性。这是 EUC 机制下采集到的用户数据具有更高可靠性的第一层保障。

b)与 BSD 相比，EUC 机制中针对用户提交的数据添加了粗糙过滤操作，针对用户所提交有明显错误的数据，直接过滤掉，并对该用户进行不可信任标记，这一过程中 MCS 平台具备自主学习的能力即会做到不断完善过的滤规则。所以一定程度上影响了前面所阐述的提交接受率(SAR)，对总的用户有效积分动态排名起到关键作用。经过粗糙过滤后的用户数据，这是可靠性的第二层保障。

c)与 BSD 相比，EUC 机制中任务分发根据地理位置进行，每个地理位置都维护一个任务列表(TL)，出现在 TL 表中的用户，表明该用户有着较高的可信度，且每次任务分发优先选取 TL 表中的用户进行任务转发。在 EUC 机制中，这是用户所上传的数据具有较高的可信度第三层保障。

由上面三点的证明，在时间与空间上的限制、数据粗糙过滤、优先选择具有历史记录上可靠性较高的用户，这些都是 BSD 机制没有考虑到的部分。可见用户上传的数据与 BSD 相比具有更高的可信度。证毕。

通过上述的分析，可以看出，本文提出的框架在在用户招募速度、转发成功率和数据可靠性均优于 BSD。在下面一节，将用实验证明本文提出机制的优点。

7　实验分析

本文实验使用 ONE 平台[23]，将本文所提出 EUC 机制与文献[8]的 Bubble-sensing 任务分发机制 BSD(Bubble-sensing distribution)进行比较，在相同条件下比较二者在任务分发过程中，其任务分发速度以及效率等。然后对本文提出冷、热采集情况进行实验，对各项指标进行比较、分析。实验中任务时间分别设置为 2、5 和 8 h 任务，用户活动范围为 2800×2800 m^2，在其中随机选取一任务监测区域，半径为 150×150 m^2，其他实验场景是一样的，情景如图 6 所示。该图为某一地区，S 为中心的圆为任务收集监测区域，有若干个用户沿着相应的轨迹移动。依赖用户的移动性、社交属性等进行任务转发，下面对两种转发机制通过实验进行性能比较。

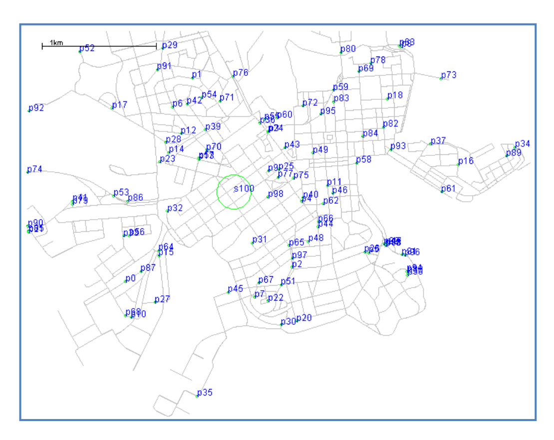


图 6　实验情景图

实验 1　初始有 2 人携带任务信息，在活动范围内有 50 用户随机移动，要求招募这 50 名用户来采集该区域监测数据，2 小时内完成。

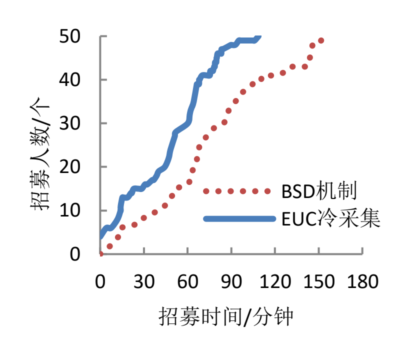


图 7　2 小时任务 BSD 与 EUC 冷采集比较

从图 7 可知，2 小时任务中，在 BSD 机制的后期，任务传播依赖用户的移动，其速度趋势变小，而对于 EUC 机制来讲，任务的传播速度十分迅速，说明用户的社交属性在任务的传播中，确实起到至关重要的作用，短时间内呈现出爆炸式的扩散。通过图 8 可知，在 EUC 热采集中，保留着一部分高效的用户，当这些高效用户收到任务后，迅速把这些任务发送给其他人，可见 EUC 热采集分发任务的速度更快。从 2 小时任务实验可以看出，EUC 机制的两种采集方式，其任务信息的传播速度明

显优于 BSD。

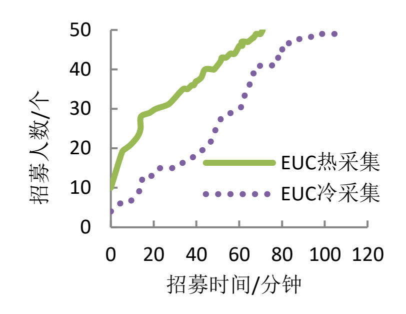


图 8 2 小时任务中 EUC 中冷、热采集比较

实验 2 初始有 5 人携带任务信息，活动范围内有 100 用户随机移动，要求至少招募 40 个用户来采集该区域监测数据，5 小时内完成。

从图 9 可知，招募用户数达到 40 个的时候，EUC 机制分发任务耗时明显少于 BSD，有大约 45%的差距。其原因是 BSD 完全依赖于用户的相遇转发，特别是在任务分发初期，扩散程度十分低，由于 EUC 机制中涉及到用户的社交属性，以及为了获得更高的奖励，用户通过社交平台，通知符合条件的好友来共同完成，从而起到任务快速传播的效果。尤其是在达到任务要求的用户数后，用户数增长几乎线性趋势。但随着时间的推移，BSD 与 EUC 几乎重合，是由于后期任务分发到达了瓶颈，携带任务的用户越来越多，很多用户已经携带了任务信息，传播相对较为困难。同时，由于时间的限制，也存在部分用户不愿意执行任务。

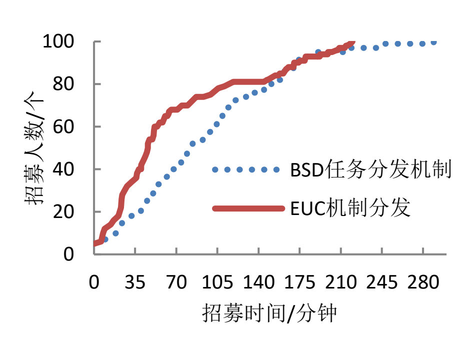


图 9 用户招募时间对比图

下面对 EUC 的冷采集与热采集两种方式进行分析。如图 10 所示，在达到任务要求的用户数时，热采集比冷采集所用时间减少了大约 33%。在冷采集过程中，由于该地理位置没有执行过感知任务，其用户列表为空，都是新用户来执行任务。而热采集时，由于该地区有过感知任务的历史记录，通过设定预定值，来选取部分历史用户直接发送任务，加快了任务分发速度。

下面讨论这三种机制招募人数的增长率。如图 11 所示，起初，每一情况都转发的较快，随着时间的推移，携带任务的用户数越来越多，任务扩散率都是从高到低，而 BSD 机制主要靠相遇转发来实现任务传播，另外两种机制相遇转发只是其中的一部分，主要靠社交属性来实现任务传播。

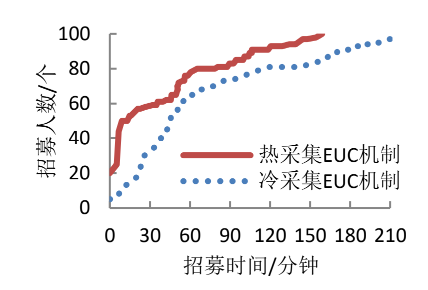


图 10 冷热采集用户招募时间对比图

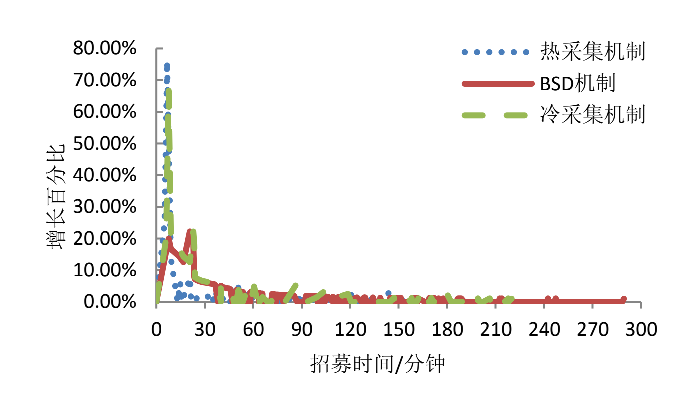


图 11 三种机制下任务扩散增长率

可见，在前 15 分钟，提出的 EUC 机制下的冷热采集，在任务传播速度上呈现极快的增长，主要是因为 EUC 机制不单单只依赖相遇转发，还涉及到了接收到任务的用户可以利用自身社交属性，实现把任务快速传播出去。在 15 分钟后可见，三种机制分发速度都出现了快速下降的趋势，尤其是 EUC 热采集机制下降最快，其次 EUC 冷采集，这两种机制下降主要原因是用户人员数量上的限制，而 BSD 机制下降的速度最为缓慢，其下降主要原因是地理位置因素的影响。综上可以看出，EUC 机制下的两种任务转发速度上相比 BSD 机制均有较大的优势。

实验 3 初始有 5 人携带任务信息，在活动范围内有 200 用户随机移动，要求招募所有用户来采集某地段监测数据，8 小时内完成。

从图 12 可知，针对较长时间的任务，任务的传播趋势基本与上述两种实验情况相同，前期用户的社交属性优势十分明显，后期由于满足条件用户的限制，传播速度明显降低。

通过三种实验情况证明结果来看，本文提出的 EUC 冷热采集与 BSD 机制采集方式可以看出，在达到相同人数的情况下，EUC 冷热采集方式，可以在短时间内快速达到要求，时间上可以得到保证。在针对实时性要求很强的情况下，相比于 BSD 机制的分发方法，本文提出的任务分发机制有着较大的优势，因为 EUC 冷热采集机制在任务开始的较短时间内，招募用户速度十分迅速，从而满足了任务的实时性要求。综上三种实验情况，本文提出的 EUC 机制在招募用户速度或人数上的有效性十分显著。

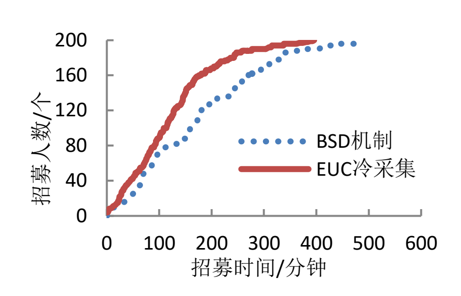


图 12　8 小时任务中 BSD 与 EUC 的冷采集比较

8　结束语

本文提出了一种基于社交属性的有效用户积分动态更新的任务分发机制，首先通过用户的社交属性，能在短时间内为任务数据采集迅速召集大量用户，完成过程中，再利用提出的两个算法，用户选择贪婪算法和有效用户排序算法，根据本文提出的用户选择贪婪算法，选出部分可信度较高用户进行任务数据采集和任务转发，然后再根据用户完成后的情况使用提出的有效用户排序算法来动态更新每个用户的有效积分，再更新到每个任务列表中。从而提高用户完成任务效率、减少任务分发时间以及降低通信开销，进而提高移动群智感知任务中任务分发机制的可用性。针对后续的研究，可以考虑用户的数据验证方面的探索。

参考文献：

[1] Stojmenovic I. Keynote 1: Mobile cloud and green computing [J]. Procedia Computer Science, 2012, 10: 18-19.

[2] 刘云浩. 群智感知计算. 中国计算机学会通信, 2012, 8 (10): 38-41.

[3] Branch J, Szymanski B, Giannella C, et al. In-network outlier detection in wireless sensor networks [J]. Knowledge & Information Systems, 2013, 34 (1): 23-54.

[4] Gungor V C, Lu B, Hancke G P. Opportunities and challenges of wireless sensor networks in smart grid [J]. IEEE Trans on Industrial Electronics, 2010, 57 (10): 3557-3564.

[5] Costa P, Coulson G, Gold R, et al. The RUNES middleware for networked embedded systems and its application in a disaster management scenario [C]// Proc of IEEE International Conference on Pervasive Computing and Communications. Washington DC: IEEE Computer Society, 2007: 69-78.

[6] Baldus H, Klabunde K, Müsch G. Reliable Set-up of medical body-sensor networks [M]// Wireless Sensor Networks. Berlin: Springer Heidelberg, 2004: 353-363.

[7] Wang Y, Huang Y, Louis C. Respecting user privacy in mobile crowdsourcing [J]. Science, 2013, 2 (2): 1–15.

[8] Lu H, Lane N D, Eisenman S B, et al. Fast track article: Bubble-sensing: binding sensing tasks to the physical world [J]. Pervasive & Mobile Computing, 2010, 6 (1): 58-71.

[9] Krontiris I, Dimitriou T. A platform for privacy protection of data requesters and data providers in mobile sensing [J]. Computer Communications, 2015, 65 (C): 43-54.

[10] Gong Y, Wei L, Guo Y, et al. Optimal Task recommendation for mobile crowdsourcing with privacy control [J]. IEEE Internet of Things Journal, 2016, 3 (5): 745-756.

[11] Zhuo G, Jia Q, Guo L, et al. Privacy-preserving verifiable data aggregation and analysis for cloud-assisted mobile crowdsourcing [C]// Proc of IEEE International Conference on Computer Communications. 2016: 1-9.

[12] Zhuo G, Jia Q, Guo L, et al. Privacy-preserving verifiable set operation in big data for cloud-assisted mobile crowdsourcing [J]. IEEE Internet of Things Journal, 2017, PP (99): 1-1.

[13] Chen Jianwei, Ma Huadong, Wei S L, et al. Participant-density-aware privacy-preserving aggregate statistics for mobile crowd-sensing [C]// Proc of IEEE International Conference on Parallel and Distributed Systems. 2015: 140-147

[14] Gisdakis S, Giannetsos T, Papadimitratos P. Security, privacy, and incentive provision for mobile crowd sensing systems [J]. IEEE Internet of Things Journal, 2016: 839-853

[15] Thepvilojanapong N, Zhang K, Tsujimori T, et al. Participation-aware incentive for active crowd sensing [C]// Proc of International Conference on High Performance Computing and Communications & IEEE International Conference on Embedded and Ubiquitous Computing. 2014: 2127-2134.

[16] Dwork C, Kenthapadi K, Mcsherry F, et al. Our data, ourselves: privacy via distributed noise generation [C]// Proc of International Conference on the Theory and Applications of Cryptographic Techniques. 2006: 486-503.

[17] Shin M, Cornelius C, Dan P, et al. AnonySense: a system for anonymous opportunistic sensing [J]. Pervasive & Mobile Computing, 2011, 7 (1): 16-30.

[18] Tathagata D, Prashanth M, Venkata N. PRISM: platform for a remote sensing using smartphone [C]// Proc of the 8th International Conference on Mobile Systems, Applications, and Services. 2010: 63-76

[19] Tuncay G S, Benincasa G, Helmy A. Autonomous and distributed recruitment and data collection framework for opportunistic sensing [J]. ACM Sigmobile Mobile Computing & Communications Review, 2012, 16 (4): 407-410.

[20] 刘媛妮，李丽. 基于社交性的移动群智感知网络自适应任务分发机 [EB/OL]. 北京：中国科技论文在线 [2016-01-18]. http://www. paper. edu. cn/releasepaper/content/201601-492.

[21] Ari K, Jorg O, Teemu K. The one simulator for DTN protocol evaluation [C]// Proc of the 2nd International Conference on Simulation Tools and Techniques. 2009.